Taxonomy and Challenges in Machine Learning-based Approaches to Detect Attacks in the Internet of Things

Omair Faraj
Internet Interdisciplinary Institute (IN3),
Universitat Oberta de Catalunya (UOC)
CYBERCAT-Center for Cybersecurity Research of Catalonia
Barcelona, Spain
ofaraj@uoc.edu

David Megías
Internet Interdisciplinary Institute (IN3),
Universitat Oberta de Catalunya (UOC)
CYBERCAT-Center for Cybersecurity Research of Catalonia
Barcelona, Spain
dmegias@uoc.edu

Abdel-Mehsen Ahmad
Lebanese International University (LIU),
The International University of Beirut (BIU)
Beirut, Lebanon
abdelmehsen.ahmad@liu.edu.lb

Joaquin Garcia-Alfaro
SAMOVAR, Telecom SudParis,
Institut Polytechnique de Paris
Évry, France
joaquin.garcia_alfaro@telecom-sudparis.eu

ABSTRACT

The insecure growth of Internet-of-Things (IoT) can threaten its promising benefits to our daily life activities. Weak designs, low computational capabilities, and faulty protocol implementations are just a few examples that explain why IoT devices are nowadays highly prone to cyber-attacks. In this survey paper, we review approaches addressing this problem. We focus on machine learning-based solutions as a representative trend in the related literature. We survey and classify Machine Learning (ML)-based techniques that are suitable for the construction of Intrusion Detection Systems (IDS) for IoT. We contribute with a detailed classification of each approach based on our own taxonomy. Open issues and research challenges are also discussed and provided.

KEYWORDS

Intrusion Detection Systems (IDS), Internet of Things (IoT), Machine Learning, Network Security

ACM Reference Format:
Omair Faraj, David Megías, Abdel-Mehsen Ahmad, and Joaquin Garcia-Alfaro. 2020. Taxonomy and Challenges in Machine Learning-based Approaches to Detect Attacks in the Internet of Things. In *The 15th International Conference on Availability, Reliability and Security (ARES 2020), August 25–28, 2020, Virtual Event, Ireland.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3407023.3407048

1 INTRODUCTION

The Internet of Things (IoT) paradigm enables physical objects with additional sensing, computing and communication capabilities [1]. It promotes the interaction of smart devices with our day life activities. From home automation, industries, and healthcare; to automotive, manufacturing, transportation and energy. IoT connects the physical world to the cyberspace [2]. These achievements rely on distributed storage, ubiquitous sensing and decentralized computing, all in a single framework [3]. In its simplest form, an IoT environment relies on sensing objects enabled with Internet protocols, services and some other networked-applications (e.g., radio-frequency identification, machine-to-machine communication and cloud computing).

IoT applications promise to improve human life. However, the insecure growth of IoT threatens such promising benefits. This has gained the attention of cybersecurity and data protection research communities. The study of security and privacy issues related to IoT is nowadays a very active area of research [4]. Concerns about the integrity, confidentiality and availability of IoT applications have been raised [5, 6]. The presence of malware, spyware and eavesdroppers may thwart IoT data and lead to major threats [7, 8]. Adequate detection methods are required to build Intrusion Detection Systems (IDS) to monitor IoT environments and avoid adversarial situations [9–11].

Executing latency-sensitive security tasks and computational-intensive ones is prohibitive to IoT devices with restricted capabilities. These devices cannot deploy existing security solutions that require a heavy computational burden, which makes them vulnerable to attacks [12]. Critical infrastructures, such as transportation and healthcare systems, and household appliances can lead to dreadful consequences when subject to attacks. This will threaten the security and privacy of families, cities and even countries. Tests were conducted on three smart home devices by Notra et al. [13], showing different vulnerabilities regarding users privacy, encryption/decryption and authentication. Traditional security approaches and countermeasures were not able to work properly due to the presence of many standards, limited computing power of devices, communication stacks and the high number of interconnected devices. These approaches may fail to defend IoT environments [14]. The computational constraints of IoT devices, their low capabilities in terms of storage, and their networking peculiarities (e.g., multi-hop transmission for the forwarding and routing of packets) are

ARES 2020, August 25–28, 2020, Virtual Event, Ireland

ACM ISBN 978-1-4503-8833-7/20/08…$15.00
https://doi.org/10.1145/3407023.3407048

also additional reasons for such failure. Specific security solutions must be developed to allow users to overcome security challenges efficiently [11]. Even with the presence of methods for enhancing the security of IoT environments, which provides confidentiality and authentication, IoT networks are still vulnerable to many attacks aiming to disrupt the network [11]. To fulfill this purpose, IDSs are proposed and designed to detect these attacks and protect IoT networks. However, the design of new detection methods to handle IoT attacks is a challenging topic [14].

An IoT network can be formed by a large number of IoT devices [15]. These devices generate an enormous quantity of data. To deal with this Big Data, Machine Learning (ML) techniques are the most dominant method. The detection of intrusions in such a huge amount of data has been extensively studied through ML and statistics. For this reason, it is very important to use intelligent tools to assist IDSs [16]. Defense policies and key parameters in the security protocols must be chosen by IoT devices intelligently when an attack occurs. These tasks are challenging and difficult to establish. ML will provide the ability to accurately estimate the current network and attack state in a real-time process [17].

In this paper, we survey recent IDS systems and methods for IoT networks based on ML. We analyze different aspects of study that should be taken into consideration during the design of an IDS for IoT. We also contribute with a taxonomy that includes different attributes related to the development of IDS for IoT. Based on our survey and taxonomy, we finally discuss some open issues and related research challenges.

Paper Organization — Section 2 provides some additional background. Section 3 explores several recent detection methods, classifies them using our own taxonomy and surveys recent work on intrusion detection for IoT. Section 4 briefly presents some open research issues and challenges. Section 5 concludes the paper.

2 BACKGROUND

This section briefly introduces key aspects of IDS systems and summarizes some related work.

2.1 Intrusion Detection

Intrusion detection is the ability to detect unusual and unauthorized actions, known as intrusions against the network. These intrusions are performed by illegitimate users, known as intruders, which can be either internal or external. Intruders aim to get an unauthorized access to the information system and misuse the data. An IDS is mainly formed from three different components: sensors attached to hosts or several network positions, an analysis entity that is responsible for detecting intrusions from the collected data sent by sensors, and a reporting system for alerting the network monitor about the detected intrusion [14, 18, 19].

Under the context of Machine Learning (ML)-based IDS for IoT protection, intrusion detection can be classified based on many aspects: placement strategy, detection method, attack type, performance evaluation, IoT scenario, machine learning technique and study methodology.

2.1.1 Placement Strategy. The IDS can be placed in different positions of the IoT network. These positions are based on the IoT network architecture that is organized into three main domains: physical, network and application. The physical domain includes the sensors that capture data from the physical environment. The network domain carries the collected data from physical domain into applications and users of the network solutions and protocols. The application domain allows users to monitor the physical objects in the physical domain by different interfaces. Based on these domains, the IDS placement strategy can be distributed, centralized or hybrid. In the distributed IDS placement, the IDS is placed in all physical objects of the network. This requires more resources in terms of processing, capacity and energy from the network nodes. Thus, the IDS must be optimized since the network nodes are resource-constrained. In the centralized IDS placement, the IDS is placed in the border router. This will detect intrusions from the internet against the physical objects in the physical domain, but it will generate communication overhead between nodes and the border router. In addition, this approach is not enough to detect attacks that involve nodes by only analyzing the traffic that traverses the border router. Finally, the hybrid IDS placement takes into account the advantages of both distributed and centralized IDS placement strategies in two main approaches: (i) arrange the network into clusters and the main node of each cluster deploys the IDS for monitoring the other nodes in its cluster, or (ii) the IDS is placed in the border router and in some dedicated nodes.

2.1.2 Detection Methods. IDSs are classified based on the detection mechanism into four different categories, namely anomaly-based, signature-based, specification-based and hybrid. In anomaly-based detection, IDSs generate an alert once the activities of a system deviate from a normal behavior profile and exceed a certain threshold. The advantage of this method is the ability to detect new attacks. However, any deviation from the normal behavior profile is considered as an attack, even if it was a legitimate activity, leading to high false-positive rates [20]. In the signature-based approach, an attack is detected by the IDS when its signature matches another one stored in the IDS database. Thus, if any activity matches patterns or signatures, stored in the database, an alert will be triggered by the reporting system. This approach is accurate for known attacks but it is not efficient for unknown ones [19, 21]. The specification-based approach, similar to the anomaly-based one, detects attacks when an activity deviates from a normal behavior. However, in this approach, there are a number of rules and thresholds that define the behavior of the network components such as nodes, routing tables, protocols and others. These rules are defined manually by an expert, which leads to a lower false-positive rate than anomaly-based approach. These solutions are time-consuming and can not be deployed in different environments with different specifications [20, 22, 23]. Hybrid approaches, on the other side, combine the main advantages of anomaly-based, signature-based and specification-based strategies.

2.1.3 Attack Types. IoT vulnerabilities include a large number of security issues. Such issues can be classified w.r.t. IoT building blocks, such as: physical objects, protocols, data and software. Data security issues can lead to practical attacks such as Denial of Service (DoS), data scavenging and brute-force attacks. Protocol issues can lead to routing attacks (e.g. sinkhole attack, selective forwarding, wormhole attack, Sybil attack, etc.) or main-in-the-middle attacks (e.g. injection of malicious code) [24].

2.1.4 Performance Evaluation. To evaluate the performance of an IDS, several metrics should be taken into consideration. These can be listed as follows:

- Accuracy rate: includes detection accuracy, classification accuracy, false positive rate, false negative rate, true positive rate, true negative rate and receiver operating characteristic curves (ROC curves).
- Complexity: amount of resources for an IDS to work efficiently (e.g. time and memory requirements).
- Scalability: capability of the system to stay efficient when security challenges, data traffic, network size or attacks increase.
- Network delay: the time for data to be delivered from one endpoint to another across the network (i.e. propagation delay, processing delay, transmission delay and queuing delay).
- Energy consumption: indicates the energy consumed by all the nodes in the network during attack detection process.
- Computational overhead: time needed by a node or a standalone detection entity to process a packet received.
- Real-time detection: real-time or offline detection and processing

2.1.5 IoT Scenario. As stated above, IoT is widely used in our daily life activities and can be deployed in almost everything around us. Therefore, IoT can be deployed in home, industries, vehicular applications, medical and health centers.

2.1.6 Machine Learning in IDS. Machine learning (ML) is a process that acts after learning from study or experience without being explicitly programmed, and thus it can improve automatically. The efficiency, reliability and cost-effectiveness that ML has brought to the computing processes have made it a major technique to be used in various fields including medical, engineering and computing [25–27]. ML relies on learning data sets taken as inputs. For this reason, it is used to enhance IoT security. ML is often applied in anomaly-based, signature-based and specification-based attack detection methods [28, 29]. ML techniques can be classified in four main categories [30]:

- supervised learning, such as regression and classification, which includes decision trees, random forest, deep learning, Bayesian, Support Vector Machines (SVM), k-Nearest Neighbor (k-NN) and Artificial Neural Networks (ANN), for which all training data are labeled;
- semi-supervised learning, which has a small amount of labeled training data;
- unsupervised learning, such as clustering —which includes k-means, hierarchical and fuzzy-c-means—, and dimensionality reduction —which includes Singular Value Decomposition (SVD), Principal Component Analysis (PCA) and Independent component analysis (ICA), for which there is no labeled training data; and, finally,
- reinforcement learning, such as Q-learning, which involves an agent, a set of states and a set of actions per state.

2.1.7 Study Methodology. Each system should be able to validate the results using a validation or study methodology. These methodologies are based on the collected data or the proposed algorithms. Study methodologies can be classified as experimental, simulation, numerical, theoretical, empirical and statistical [31]. These are used as an attribute for the IDS study.

2.2 Related Work

Some reviews have been conducted regarding intrusion detection in the fields of cloud computing, Wireless Sensor Networks (WSN) and traditional networks [32–34]. However, only a few surveys are focused on intrusion detection methods in IoT environments. In this section, we list some recent surveys and review papers that discuss intrusion detection in IoT networks. These works are used to build the taxonomy proposed in this paper and indicate the missing aspects that researchers must take into consideration while developing a new IDS for IoT.

In [35], Hajiheidari et al. present a systematic literature review of the IDSs in IoT categorizing these systems based on detection method, placement strategy and some specific attack types. They discuss the advantages and disadvantages of the selected approaches and provide some future trends in this field of study. The survey does not take into consideration the ML techniques used in the selected mechanisms and focus on some attacks only.

da Costa et al. [36] review intrusion detection approaches based on ML techniques. The work studies the approaches in terms of communication protocol, application protocol, data format, ML technique and precision rate. The authors do not take into account many aspects of the classification as security threats, IoT scenarios, performance metrics rather than precision rate, detection method and placement strategy.

Zarpelão et al. [37] present a survey of research efforts in intrusion detection for IoT environments. They classify IDSs based on the following attributes: detection method, placement strategy, validation strategy and security threat. Nevertheless, performance evaluation metrics and ML techniques are not present in their classification of the selected research efforts.

In [38], Benkhelifa et al. present advancements in IDSs for IoT technologies, focusing on the type of network architecture. The review classifies the surveyed IDSs based on detection method, network architecture, communication technology, IDS type (i.e. Host-based IDS (HIDS) or Network-based IDS (NIDS)). The authors provide future directions for IDSs in IoT. Similar to other mentioned papers, many IDS attributes are not discussed.

Tabassum et al. [39] present a survey on recent approaches in IDS for IoT. The focus of their survey is on IDSs that use hybrid and intelligent techniques. The authors provide a review on IoT protocols, layers and their security issues ending up with future directions for the implementation of IDS based on the limitations and advantages of the selected approaches. This survey concentrates on the classification and categorization of IDS approaches on: IoT architecture, IoT communication protocols, detection methods, placement strategy and ML techniques.

Gendreau and Moorman [40], provide a survey of IDS techniques in IoT. They present a brief summary on IDS to understand the IDS platform. Then, they discuss some selected IDSs used in IoT environments classifying these systems based on detection technique, features and interaction ability score. Some guidelines for the deployment of IDSs and open research problems are outlined.

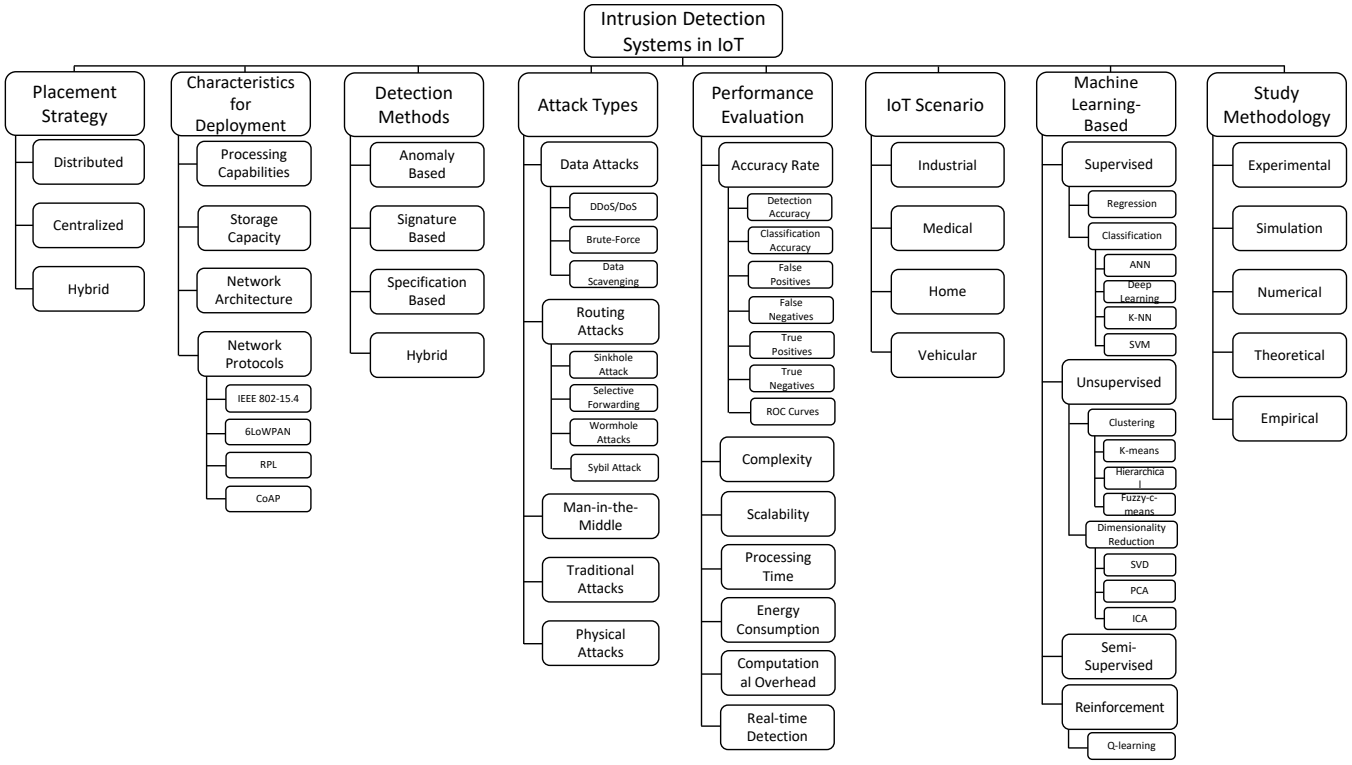

Figure 1: Taxonomy of Intrusion Detection Systems for IoT.

The survey does not take into consideration most of the attributes needed to study IDS for IoT approaches.

Santos et al. [41] present a survey on a number of selected IDSs for IoT networks to study the development of intrusion detection. They classify the works based on: detection method, security thread and placement strategy. The survey does not take into account many attributes needed to study the development of IDS in IoT networks.

3 INTRUSION DETECTION FOR IOT

In this section, we present some relevant IDS approaches for different IoT environments. The approaches are then classified in Tables 1 and 2, where "-" denotes that the attribute is not specified. Classification is based on the attributes mentioned in Section 2.1. Figure 1 provides our proposed taxonomy.

In [42], an offline IDS is proposed for an IoT network based on Artificial Neural Networks (ANN) against denial-of-service attacks. The focus of the work is on classifying the received patterns as normal or malicious activity. A feed-forward learning algorithm and a backward learning algorithm are used to train the network.

In [43], authors propose an intrusion detection and mitigation framework implemented by using a Software-Defined Networking (SDN) architecture and its enabling communication protocol, OpenFlow, in the Philips Hue lighting system as a smart-home environment. They performed commands using the Hue App such as: switching On/Off, color setting and brightness. Then an attack script is launched from a machine in the same network. Lastly, they evaluated the results of the IDS installed on a centralized SDN controller.

Detection of potential attacks in smart homes, in particular at the physical layer, by detecting deviations from legitimate communication behavior is presented in [44]. The method is based on analyzing the Radio Signal Strength Indication (RSSI) which is associated to the connected devices in the network. Suspicious wireless transmissions are characterized using a ML neural network algorithm.

A game-theoretic approach for anomaly detection in low-resource IoT devices is proposed in [45]. The concept of game theory with the help of Nash equilibrium is used to predict the state that allows the devices to activate an anomaly detection technique to detect new attacks. The method balances between energy consumption and accuracy detection by activating the IDS in each device only when a new attack is expected to occur.

In [46], authors propose an intrusion detection model using genetic programming with K-Nearest Neighbor classifier against DoS, Probing (PRB), Remote to User (R2L) and User to Root (U2R) attacks. They use the KDD CUP[1] data set [58] that was fed into a pre-processing stage to find the optimal features that will be used in the classification process. The classification process carried out by a k-NN classifier determines whether there is suspicious data.

InDReS is an Intrusion Detection and Response System developed to detect sinkhole attacks in an IoT network with 6LoWPAN [47]. In this system, nodes are isolated and network is reconstructed after an attack is detected. The work is compared with

[1] A database which contains a standard set of data. This data can be audited. It includes a large number of intrusions that are simulated in a military network. It was one of the most widely used data sets for the evaluation of anomaly detection systems [58].

Table 1: Recent intrusion detection systems for IoT networks.

Reference	Method	Placement Strategy	Detection Method	Attack Type	IoT Scenario	Machine Learning	Study Methodology
[42]	Classifying normal and threat patterns in an IoT network using ML	Centralized	Anomaly-based	DDoS/DoS	-	NN	Simulation
[43]	Detecting suspicious activities in home devices using OpenFlow	Centralized	Signature-based	Routing attacks, man-in-the-middle	Home	Regression, SVM	Experiment
[44]	IDS approach based on radio communication profiling and monitoring using ML	Centralized	Anomaly-based	Physical attacks, unauthorized wireless transmissions	Home	NN	Experiment
[45]	Detection technique that balances between energy consumption and detection accuracy using game theory	Hybrid	Anomaly-based	DoS	WSN	NN	Simulation
[46]	Intrusion detection model based on genetic programming and K-NN for classifying data as normal or suspicious	Centralized	Anomaly-based	DoS, R2L, U2R, PRB	-	K-NN	Simulation
[47]	Detect sinkhole attacks relying on constraint based specification model	Hybrid	Specification-based	Sinkhole	-	-	Simulation
[48]	IDS based on Deep Belief Network (DBN) and Genetic Programming (GP) to generate optimal network structure	Centralized	Anomaly-based	DoS, R2L, PRB, U2R	-	Deep Learning	Simulation
[49]	Localizing malicious nodes and detecting intrusions for WSN and gateways using ML	Hybrid	Anomaly-based	Selective forward, conventional IP attacks	-	SVM and Deep Learning	Simulation
[50]	Detecting network cyber attacks using supervised three layer IDS	Centralized	Anomaly-based	DoS, man-in-the-middle, spoofing, reconnaissance, replay	Home	Supervised	Empirical
[51]	Intrusion detection and prevention system using Genetic Programming (GP)	Centralized	Anomaly-based	DoS, Probe, R2L, U2R	-	-	Simulation
[52]	Threat detection for IoT using supervised learning algorithm to solve authentication issues	Centralized	Signature-based	-	-	ANN	Simulation
[53]	A Bi-directional Long Short-Term Memory Recurrent Neural Network (BLSTM RNN) for detecting attacks in IoT networks	Centralized	Anomaly-based	Analysis, Backdoor, DoS, Worms, reconnaissance	-	Deep learning	Simulation
[54]	Detecting intrusions against routing protocol (RPL) attacks using genetic programming	Centralized	Signature-based	RPL (hello flood, version number attacks)	Industrial	-	Simulation
[55]	Extension of SVELTE IDS using Expected Transmissions (ETX) metric	Hybrid	Anomaly-based	Selective forward attack	6LoWPAN-network		Simulation
[56]	Anomaly detection framework focusing on IoT specific features applied to IoT botnets	Hybrid	Anomaly-based	DDoS	Home	K-NN, SVM, DT, RF and NN	Simulation
[57]	IDS based on SFC and PCA	Hybrid	Anomaly-based	-	-	SFC, PCA	Simulation

Table 2: Performance evaluation of IDSs in IoT

Reference	Detection Accuracy	Classification Accuracy	TPR	FPR	TNR	FNR	ROC curves	Processing time	Energy consumption	Computation overhead	Real-time detection
[42]	-	99%	99.4%	0.6%	-	-	-	-	-	-	Offline
[43]	94.25%	85.05%	35.47%	5.74%	-	-	-	-	-	-	Real-time
[44]	-	-	-	-	-	-	-	-	-	-	-
[45]	90-98%	-	-	-	-	-	-	900s including simulation	2-3J	-	Real-time
[46]	99%	-	-	-	-	-	-	-	-	-	Real-time
[47]	-	-	-	-	-	-	-	-	0.06-0.1J (Avg)	-	Real-time
[48]	61-99%	97-99%	-	-	-	-	-	-	-	-	Real-time
[49]	95%	-	-	-	-	-	-	50-600s execution time	-	-	Real-time
[50]	90-98%	-	89-99%	-	-	-	-	0.1-0.4s classification time	-	-	Real-time
[51]	78-81%	-	-	-	-	-	-	-	-	-	Real-time
[52]	84%	-	-	8%	-	-	-	-	-	-	Real-time
[53]	95%	-	96%	0.23%	0.0237%	3.94%	-	2.19s classification time	-	-	Real-time
[54]	96-99%	-	95-99%	0.2-1.1%	-	-	-	0.5-5s detection time	-	-	Real-time
[55]	-	-	90-98%	-	-	-	-	-	-	Power: 1.5-22mW per node, RAM: 5570 Bytes and ROM: 6 Bytes	Real-time
[56]	99%	91-99%	-	-	-	-	-	-	-	-	Real-time
[57]	80-97.4%	-	-	1.5%	-	-	-	0.1-0.6s detection time	-	-	Real-time

the INTI (Intrusion detection of SiNkhole attacks on 6LoWPAN for InterneT of ThIngs) scheme [59] regarding throughput, packet drop ratio and energy consumption.

An intrusion detection model based on Genetic Programming (GP) and Deep Belief Network (DBN) is proposed in [48]. It is used to detect low-frequency attacks by reducing the network complexity through generating the optimal number of layers and neurons in the network reaching an optimal network structure.

A hierarchical intrusion detection algorithm for the WSNs and IoT gateways that connect them to the internet is proposed in [49]. The detection system uses support vector machines for WSN intrusion detection and deep learning for IoT gateway intrusion detection. It localizes malicious nodes by the combination of machine learning classification and a statistical approach to reach a trade-off between detection efficiency and resource overhead.

A three-layer IDS for smart home IoT devices is proposed in [50]. It is used to detect malicious network activities and identify the type of attack deployed on any device in the network. The system mainly consists of three functions: setting the normal behavior of each connected device in the network, identifying malicious packets during an attack to the network and classifying the type of an attack based on four network attack categories.

Authors in [51], propose an intrusion detection and prevention system using Genetic Programming to detect suspicious traffic. The system has a blocking mechanism once malicious activity is detected from an IoT device. The work is based on SDN technology as a controller to handle large data traffic.

In [52], authors consider a threat detection system in an IoT network to solve authentication issues. The system uses the Artificial Neural Networks (ANN) machine learning approach to detect attacks and discard data after classifying it as a threat. They used the UNSW-15[2] and the NSL-KDD[3] data sets for training and testing their algorithm.

A deep learning technique is proposed by [53] for detecting attacks in IoT networks. The method uses Bi-directional Long Short-Term Memory Recurrent Neural Network (BLSTM RNN) focusing on binary classification of normal and suspicious patterns in the network.

Genetic programming is used in [54] to detect routing attacks in IPv6 Routing Protocol for Low-Power and Lossy Networks (RPL) in an industrial IoT network. The work proposes a methodology that aims to answer the following three research questions: Can genetic programming be used to detect intrusion in industrial IoT (IIoT)?, what is the suitable IDS for IIoT? and how well this system performs in terms of detection accuracy and detection time? After simulation, results show successful deployment of genetic programming in the detection of routing attacks in IIoT. They obtain high accuracy, high true positive and low false positive rates.

Shreenivas et al. [55], extend SVELTE IDS [62], an IDS for IoT, with an intrusion detection module using ETX (Expected Transmissions) metric. The proposed methodology attempts to detect attacks against RPL routing protocol, used in 6LoWPAN networks. It also identifies malicious nodes that conduct attacks against the network by using geographic hints.

[2]Data set created by the establishment of the synthetic environment at the UNSW cybersecurity lab. It represents nine major families of attacks [60]

[3]Improved version of KDD'99 data set. A data set suggested to overcome the problems found in KDD'99 [61].

An intrusion detection technique is proposed in [56] for DDoS detection in IoT networks. It uses IoT-specific network behavior for feature selection with several machine learning algorithms to detect IoT device sources of DDoS attacks. K-NN, SVM, Decision Tree (DT), Random Forest (RF) and NN machine learning techniques are used and detection accuracy was compared among these techniques.

To improve effectiveness of intrusion detection, a ML method is proposed in [57]. It uses suppressed fuzzy clustering (SFC) and principal component analysis (PCA). The method classifies data as high risk and low risk data based on high frequency and low frequency measures. The algorithm is compared with the neural network algorithm and Bayesian algorithm obtaining an improved applicability and better adaptability.

4 OPEN ISSUES AND RESEARCH CHALLENGES

After classifying the research efforts in IDS for IoT networks in the previous section, we observed that the current systems still need more research to investigate the weak points of each method and, also, to combine as many aspects of study as possible in the research work. Addressing security challenges and vulnerabilities in IoT architectures should be taken into account by researchers. For this reason, in this section, we address the main issues that should be taken into consideration when developing a new IDS or when proposing a solution for intrusion detection in IoT networks. The identified open issues and research gaps are detailed below.

4.1 Limitations of surveyed solutions

Limitations inferred from the previously surveyed solutions can be categorized into four main groups: typical aspects, attack detection, emerging technologies and performance analysis.

4.1.1 Typical aspects: It is required to carry out a detailed study on the advantages and disadvantages of the previously used aspects: placement strategy, detection method and machine learning techniques. This kind of study is still missing in most IDS implementations. There is no clear reference to know the advantages and disadvantages of each of the mentioned aspects that are essential for the development of a new IDS.

Additionally, the study methodology, which is one of the most important components when proposing a solution, must be improved. It is not enough to study the proposed algorithm using only one studying methodology. Simulation and experimental methods must be addressed and results from different methods should be compared, such that IDSs can be evaluated on real network traces.

4.1.2 Attack detection: More attack types must be covered, and IDSs must be proposed for a wide range of attack types rather than focusing on known ones. In the evolution of IoT networks and their applications, many new types of attacks are being deployed by malicious parties and eavesdroppers to have unauthorized access to the network. Such attacks need to be introduced to the research efforts in designing and implementing security solutions rather than just concentrating on the traditional and known attacks. In the process of detecting attacks, it is very important to reach the minimum processing time for the procedure to be completed. In real-time detection and delay-sensitive services, researchers need to establish algorithms and procedures that require very small

processing time. This will also help to reduce the consumed energy and power in the network.

4.1.3 Emerging technologies: IDSs must be proposed for different IoT technologies. Many researchers and organizations have proposed communication technologies and standards for IoT applications [63]. These technologies are used in routing, communication and integration between the internet and the network. The most popular technologies used in IoT networks are: IEEE802.15.4 for physical and medium access control layers, Bluetooth Low Energy (BLE) —an evolution of Bluetooth technology for low power devices—, WirelessHART for industrial process control, Z-wave for automation of small businesses and homes, the RPL routing protocol, 6LoWPAN standard to adapt the IPv6 packet to IEEE802.15.4, and CoAP and MQTT protocols for the application layer. These technologies are one of the main characteristics to be explored by IDSs to develop a system for the detection of security threats. New technologies, such as CoAP, BLE, Z-Wave and WirelessHART, which are commonly found in the market, need to be addressed and studied by IDS solutions rather than concentrating efforts only on previous technologies in IoT networks.

4.1.4 Performance analysis: Energy and power of network nodes must be studied along with detection accuracy and processing time. These two attributes are very important in IoT networks, since IoT devices may have low energy and power capabilities and most of the security solutions require more energy and power than the devices can hold. This opens the challenge of lightweight solutions to be adopted by researchers. Moreover, researchers should take into account the challenges in security for IoT networks, such as scalability, hardware limitations of nodes, services that are delay-sensitive and the interaction between all layers of the IoT network architecture. Delay-sensitive services are critical and should be extensively studied. Any security problem in such services leads to dangerous consequences, especially when these services are found in medical, military and home appliance applications. Finally, one of the important tools to be used in decision making and especially in detecting attacks is the receiver operating characteristic curve (ROC). Table 2 shows that none of reviewed systems selected in our work from recent published papers used ROC curves in their study. ROC curves compare the two operating characteristics: True Positive Rate (TPR) and False Positive Rate (FPR). This is used in the process of classifying whether we have attack or not.

4.2 Further lines for research

Most of the surveys mentioned in Section 2.2 overlook many aspects that are needed for studying an IDS in IoT networks. These attributes are very important to specify the weak points in such systems. This section discusses the limitations in the recent methods developed to detect attacks in IoT networks. This opens a great opportunity for researchers to find solutions for such limitations and problems in the recent approaches. Therefore, all the aspects mentioned in the proposed taxonomy and through the sections above need to be included in the researchers efforts to overcome IoT network vulnerabilities. These aspects are a must for the classification, categorization, improvement and analysis for the new developed methods.

In signature detection, predefined attack patterns need to be matched with the current behavior of the network. These signatures are stored on the device and each signature matches a specific attack. Generally, signature detection methods are simple to use. However, they need a signature for each attack and should store this signature on devices. This requires storage capabilities and knowledge of each attack. These requirements grow as the number of attacks increases. Anomaly detection techniques determines the ordinary behavior and uses it as a baseline to detect anomalies in the network. Deviations from this behavior is considered as an attack. These techniques are able to detect any attack and can be deployed in different environments, but it has high false positive rates and high false negative rates. Since some deviations from the normal behavior may be ordinary and a number of attacks may have a small deviation which is considered within the baseline [62].

Moreover, the enormous quantity of data generated in IoT networks lead to the need of intelligent tools to assist IDSs. These tools are established using Machine Learning techniques. Nowadays, IDSs are developing rapidly with the presence of these techniques. On the other hand, the robustness of these systems becomes questionable in the presence of adversarial attacks. A new arising framework is being developed known as Generative Adversarial Network (GAN) used to evade and deceive any IDS. It is used to fool machine learning algorithms. For this reason, the development of GAN should be used by researchers to improve recent IDSs and establish new ones to reach a system that can not be broken by robust attacks.

A final solution to handle attack detection is the use of challenge-response mechanisms, e.g., via watermarking authentication techniques. Such techniques are popular for the protection of wireless sensor networks, specially in the context of IoT environments [64, 65]. From all the limitations and restrictions mentioned before, watermarking can be used to implement anomaly detection [66]. This is being tested in a number of new projects that consider watermarking as a solution for secure transmission, data integrity, authentication and confidentiality in cyber-physical systems. Watermarking techniques are attractive since they can be deployed with lightweight calculations and less energy consumption. Hence these solutions are power efficient, can be applied for large networks with no additional overhead on network communication and storage capacity of nodes, and reduce end-to-end delay [67–73].

5 CONCLUSION

We have reviewed recent IDSs using Machine Learning approaches for IoT networks. First, the main concepts of IDS were addressed. Then, we provided a detailed review and classification of 16 selected papers (published between 2016 and 2019) for intrusion detection based on a proposed taxonomy. Finally, we have identified open issues and research challenges to improve these IDS schemes. We have observed that IDSs need to study detection rates, false positive rates, real-time detection, computation overhead and energy consumption in a combined manner. Researchers must consider all these aspects while designing and implementing a new IDS. In addition, more research should be conducted to cover all attack types and recent IoT technologies. Furthermore, research efforts are needed to find the optimal placement strategies to compute machine learning-based detection that could benefit to the security of IoT networks, while minimizing the risk of increasing the attack surface. It is also clear that anomaly detection requires building a

normal behavior and suffer from high false positive rates. Signature-based methods are considered to be easier to deploy, but require storage capabilities and knowledge of each attack. For this reason, watermarking algorithms are recommended to be deployed that are much lighter and require less power, storage and computational capabilities. Our future research will explore this research direction.

ACKNOWLEDGMENTS

The authors graciously acknowledge support from the European Commission, under grant agreement 830892 (H2020 SPARTA project), Spanish Ministry of Science and Innovation, through grant RTI2018-095094-B-C22 "CONSENT" and Agency for Management of University and Research Grants (AGAUR), through grant 2020 FI_B 00863.

REFERENCES

[1] Y. Yang, L. Wu, G. Yin, L. Li, and H. Zhao. A survey on security and privacy issues in internet-of-things. *IEEE Internet of Things Journal*, 4(5):1250–1258, 2017. ISSN 2372-2541. doi: 10.1109/JIOT.2017.2694844.

[2] Daniele Miorandi, Sabrina Sicari, Francesco De Pellegrini, and Imrich Chlamtac. Internet of things: Vision, applications and research challenges. *Ad Hoc Networks*, 10(7):1497 – 1516, 2012. ISSN 1570-8705. doi: https://doi.org/10.1016/j.adhoc.2012.02.016.

[3] Biljana L. Risteska Stojkoska and Kire V. Trivodaliev. A review of internet of things for smart home: Challenges and solutions. *Journal of Cleaner Production*, 140:1454 – 1464, 2017. ISSN 0959-6526. doi: https://doi.org/10.1016/j.jclepro.2016.10.006.

[4] J. Du and S. Chao. A study of information security for m2m of iot. In *2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE)*, volume 3, pages V3–576–V3–579, 2010. doi: 10.1109/ICACTE.2010.5579563.

[5] J. Lee and H. Kim. Security and privacy challenges in the internet of things [security and privacy matters]. *IEEE Consumer Electronics Magazine*, 6(3):134–136, 2017. ISSN 2162-2256. doi: 10.1109/MCE.2017.2685019.

[6] M. Plachkinova and C. Maurer. Teaching case security breach at target. *Journal of Information Systems Education*, 29(1):11–20, 2018. ISSN 10553096.

[7] Sergio Castillo-Perez and Joaquin Garcia-Alfaro. Spyware-based Menaces Against Web Applications. In *International Conference on Intelligent Networking and Collaborative Systems (INCOS'09)*, pages 409–412. IEEE, November 2009. doi: 10.1109/INCOS.2009.31. URL http://dx.doi.org/10.1109/INCOS.2009.31.

[8] Nizar Kheir, Gregory Blanc, Hervé Debar, Joaquin Garcia-Alfaro, and Dingqi Yang. Automated classification of C & C connections through malware URL clustering. In *IFIP International Information Security Conference*, pages 252–266. Springer, 2015. doi: 10.1007/978-3-319-18467-8_17. URL http://dx.doi.org/10.1007/978-3-319-18467-8_17.

[9] Qi Jing, Athanasios V. Vasilakos, Jiafu Wan, Jingwei Lu, and Dechao Qiu. Security of the internet of things: perspectives and challenges. *Wireless Networks*, 20(8):2481–2501, 2014. ISSN 1572-8196. doi: 10.1007/s11276-014-0761-7.

[10] H. Ning, H. Liu, and L. T. Yang. Cyberentity security in the internet of things. *Computer*, 46(4):46–53, 2013. ISSN 1558-0814. doi: 10.1109/MC.2013.74.

[11] S. Sicari, A. Rizzardi, L.A. Grieco, and A. Coen-Porisini. Security, privacy and trust in internet of things: The road ahead. *Computer Networks*, 76:146 – 164, 2015. ISSN 1389-1286. doi: https://doi.org/10.1016/j.comnet.2014.11.008.

[12] Jun Zhou, Zhenfu Cao, Xiaolei Dong, and Athanasios V. Vasilakos. Security and privacy for cloud-based iot: Challenges. *Comm. Mag.*, 55(1):26–33, 2017. ISSN 0163-6804. doi: 10.1109/MCOM.2017.1600363CM.

[13] Sukhvir Notra, Muhammad Siddiqi, Hassan Habibi Gharakheili, Vijay Sivaraman, and Roksana Boreli. An experimental study of security and privacy risks with emerging household appliances. *2014 IEEE Conference on Communications and Network Security*, pages 79–84, 2014.

[14] Bruno Bogaz Zarpelão, Rodrigo Sanches Miani, Cláudio Toshio Kawakani, and Sean Carlisto de Alvarenga. A survey of intrusion detection in internet of things. *Journal of Network and Computer Applications*, 84:25 – 37, 2017. ISSN 1084-8045. doi: https://doi.org/10.1016/j.jnca.2017.02.009.

[15] Mohiuddin Ahmed, Abdun Naser Mahmood, and Jiankun Hu. A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, 60:19 – 31, 2016. ISSN 1084-8045. doi: https://doi.org/10.1016/j.jnca.2015.11.016.

[16] Kelton A.P. Costa, Luis A.M. Pereira, Rodrigo Y.M. Nakamura, Clayton R. Pereira, João P. Papa, and Alexandre Xavier Falcão. A nature-inspired approach to speed up optimum-path forest clustering and its application to intrusion detection in computer networks. *Information Sciences*, 294:95 – 108, 2015. ISSN 0020-0255. doi: https://doi.org/10.1016/j.ins.2014.09.025.

[17] Liang Xiao, Xiaoyue Wan, Xiaozhen Lu, Yanyong Zhang, and Di Wu. Iot security techniques based on machine learning. *CoRR*, abs/1801.06275, 2018. URL http://arxiv.org/abs/1801.06275.

[18] Qais Saif Qassim, Norziana Jamil, Maslina Daud, Ahmed Patel, and Norhamadi Ja'affar. A survey of intrusion detection and prevention systems. *Information Management & Computer Security*, 18(4):277 – 290, 2010. ISSN 0968-5227. doi: https://doi.org/10.1108/09685221011079199.

[19] John R. Vacca. *Computer and Information Security Handbook*. Morgan Kaufmann, 2013. ISBN 978-0-12-394397-2.

[20] Robert Mitchell and Ing-Ray Chen. A survey of intrusion detection techniques for cyber-physical systems. *ACM Computing Surveys (CSUR)*, 46(4):55:1–55:29, 2014. ISSN 0360-0300. doi: http://doi.acm.org/10.1145/2542049.

[21] Hung-Jen Liao, Chun-Hung Richard Lin, Ying-Chih Lin, and Kuang-Yuan Tung. Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications*, 36(1):16 – 24, 2013. ISSN 1084-8045. doi: https://doi.org/10.1016/j.jnca.2012.09.004.

[22] J. P. Amaral, L. M. Oliveira, J. J. P. C. Rodrigues, G. Han, and L. Shu. Policy and network-based intrusion detection system for ipv6-enabled wireless sensor networks. In *2014 IEEE International Conference on Communications (ICC)*, pages 1796–1801, 2014. doi: 10.1109/ICC.2014.6883583.

[23] I. Butun, S. D. Morgera, and R. Sankar. A survey of intrusion detection systems in wireless sensor networks. *IEEE Communications Surveys Tutorials*, 16(1):266–282, 2014. ISSN 2373-745X. doi: 10.1109/SURV.2013.050113.00191.

[24] Hezam Akram Abdul-Ghani, Dimitri Konstantas, and Mohammed Mahyoub. A comprehensive iot attacks survey based on a building-blocked reference model. *International Journal of Advanced Computer Science and Applications*, 9(3), 2018. doi: 10.14569/IJACSA.2018.090349.

[25] Tom M. Mitchell. *Machine Learning*. McGraw Hill, 1997. ISBN 978007042807200704280 77.

[26] Taiwo Oladipupo Ayodele. *New Advances in Machine Learning*. InTech, 2010. ISBN 978-953-307-034-6.

[27] P. Langley and H.A. Simon. Applications of machine learning and rule induction. *Communications of the ACM*, 38(11):54–64, 1995. doi: 10.1145/219717.219768.

[28] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac, and P. Faruki. Network intrusion detection for iot security based on learning techniques. *IEEE Communications Surveys Tutorials*, 21(3):2671–2701, 2019. ISSN 2373-745X. doi: 10.1109/COMST.2019.2896380.

[29] C. Kolias, G. Kambourakis, A. Stavrou, and S. Gritzalis. Intrusion detection in 802.11 networks: Empirical evaluation of threats and a public dataset. *IEEE Communications Surveys Tutorials*, 18(1):184–208, 2016. ISSN 2373-745X. doi: 10.1109/COMST.2015.2402161.

[30] D. Praveen Kumar, Tarachand Amgoth, and Chandra Sekhara Rao Annavarapu. Machine learning algorithms for wireless sensor networks: A survey. *Information Fusion*, 49:1 – 25, 2019. ISSN 1566-2535. doi: https://doi.org/10.1016/j.inffus.2018.09.013.

[31] Vilhelm Verendel. Quantified security is a weak hypothesis: A critical survey of results and assumptions. In *Proceedings of the 2009 Workshop on New Security Paradigms Workshop*, pages 37–50, 2009. ISBN 978-1-60558-845-2. doi: 10.1145/1719030.1719036.

[32] Behrouz Pourghebleh and Nima Jafari Navimipour. Towards efficient data collection mechanisms in the vehicular ad hoc networks. *International Journal of Communication Systems*, 32(5):e3893, 2019. doi: 10.1002/dac.3893.

[33] Bahram Hajimirzaei and Nima Jafari Navimipour. Intrusion detection for cloud computing using neural networks and artificial bee colony optimization algorithm. *ICT Express*, 5(1):56 – 59, 2019. ISSN 2405-9595. doi: https://doi.org/10.1016/j.icte.2018.01.014.

[34] Yalda Ebadi and Nima Jafari Navimipour. An energy-aware method for data replication in the cloud environments using a tabu search and particle swarm optimization algorithm. *Concurrency and Computation: Practice and Experience*, 31(1):e4757, 2019. doi: 10.1002/cpe.4757.

[35] Somayye Hajiheidari, Karzan Wakil, Maryam Badri, and Nima Jafari Navimipour. Intrusion detection systems in the internet of things: A comprehensive investigation. *Computer Networks*, 160:165 – 191, 2019. ISSN 1389-1286. doi: https://doi.org/10.1016/j.comnet.2019.05.014.

[36] Kelton A.P. da Costa, João P. Papa, Celso O. Lisboa, Roberto Munoz, and Victor Hugo C. de Albuquerque. Internet of things: A survey on machine learning-based intrusion detection approaches. *Computer Networks*, 151:147 – 157, 2019. ISSN 1389-1286. doi: https://doi.org/10.1016/j.comnet.2019.01.023.

[37] Bruno Bogaz Zarpelão, Rodrigo Sanches Miani, Cláudio Toshio Kawakani, and Sean Carlisto de Alvarenga. A survey of intrusion detection in internet of things. *Journal of Network and Computer Applications*, 84:25 – 37, 2017. ISSN 1084-8045. doi: https://doi.org/10.1016/j.jnca.2017.02.009.

[38] E. Benkhelifa, T. Welsh, and W. Hamouda. A critical review of practices and challenges in intrusion detection systems for iot: Toward universal and resilient systems. *IEEE Communications Surveys Tutorials*, 20(4):3496–3509, 2018. ISSN 2373-745X. doi: 10.1109/COMST.2018.2844742.

[39] A. Tabassum, A. Erbad, and M. Guizani. A survey on recent approaches in intrusion detection system in iots. In *2019 15th International Wireless Communications Mobile Computing Conference (IWCMC)*, pages 1190–1197, 2019. doi: 10.1109/IWCMC.2019.8766455.

[40] A. A. Gendreau and M. Moorman. Survey of intrusion detection systems towards an end to end secure internet of things. In *2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud)*, pages 84–90, 2016. doi: 10.1109/FiCloud.2016.20.

[41] Leonel Santos, Carlos Rabadão, and Ramiro Gonçalves. Intrusion detection systems in internet of things: A literature review. In *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)*, pages 1–7, 2018. doi: 10.23919/CISTI.2018.8399291.

[42] E. Hodo, X. Bellekens, A. Hamilton, P. Dubouilh, E. Iorkyase, C. Tachtatzis, and R. Atkinson. Threat analysis of iot networks using artificial neural network intrusion detection system. In *2016 International Symposium on Networks, Computers and Communications (ISNCC)*, pages 1–6, 2016. doi: 10.1109/ISNCC.2016.7746067.

[43] M. Nobakht, V. Sivaraman, and R. Boreli. A host-based intrusion detection and mitigation framework for smart home iot using openflow. In *2016 11th International Conference on Availability, Reliability and Security (ARES)*, pages 147–156, 2016. doi: 10.1109/ARES.2016.64.

[44] J. Roux, É. Alata, G. Auriol, V. Nicomette, and M. Kâaniche. Toward an intrusion detection approach for iot based on radio communications profiling. In *2017 13th European Dependable Computing Conference (EDCC)*, pages 147–150, 2017. doi: 10.1109/EDCC.2017.11.

[45] H. Sedjelmaci, S. M. Senouci, and M. Al-Bahri. A lightweight anomaly detection technique for low-resource iot devices: A game-theoretic methodology. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6, 2016. doi: 10.1109/ICC.2016.7510811.

[46] S. Malhotra, V. Bali, and K. K. Paliwal. Genetic programming and k-nearest neighbour classifier based intrusion detection model. In *2017 7th International Conference on Cloud Computing, Data Science Engineering - Confluence*, pages 42–46, 2017. doi: 10.1109/CONFLUENCE.2017.7943121.

[47] M. Surendar and A. Umamakeswari. Indres: An intrusion detection and response system for internet of things with 6lowpan. In *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pages 1903–1908, 2016. doi: 10.1109/WiSPNET.2016.7566473.

[48] P. Li and Y. Zhang. A novel intrusion detection method for internet of things. In *2019 Chinese Control And Decision Conference (CCDC)*, pages 4761–4765, 2019. doi: 10.1109/CCDC.2019.8832753.

[49] A. Yahyaoui, T. Abdellatif, and R. Attia. Hierarchical anomaly based intrusion detection and localization in iot. In *2019 15th International Wireless Communications Mobile Computing Conference (IWCMC)*, pages 108–113, 2019. doi: 10.1109/IWCMC.2019.8766574.

[50] E. Anthi, L. Williams, M. Słowińska, G. Theodorakopoulos, and P. Burnap. A supervised intrusion detection system for smart home iot devices. *IEEE Internet of Things Journal*, 6(5):9042–9053, 2019. ISSN 2372-2541. doi: 10.1109/JIOT.2019.2926365.

[51] A. Mansour, M. Azab, M. R. M. Rizk, and M. Abdelazim. Biologically-inspired sdn-based intrusion detection and prevention mechanism for heterogeneous iot networks. In *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pages 1120–1125, 2018.

[52] S. Hanif, T. Ilyas, and M. Zeeshan. Intrusion detection in iot using artificial neural networks on unsw-15 dataset. In *2019 IEEE 16th International Conference on Smart Cities: Improving Quality of Life Using ICT IoT and AI (HONET-ICT)*, pages 152–156, 2019. doi: 10.1109/HONET.2019.8908122.

[53] B. Roy and H. Cheung. A deep learning approach for intrusion detection in internet of things using bi-directional long short-term memory recurrent neural network. In *2018 28th International Telecommunication Networks and Applications Conference (ITNAC)*, pages 1–6, 2018. doi: 10.1109/ATNAC.2018.8615294.

[54] E. Aydogan, S. Yilmaz, S. Sen, I. Butun, S. Forsström, and M. Gidlund. A central intrusion detection system for rpl-based industrial internet of things. In *2019 15th IEEE International Workshop on Factory Communication Systems (WFCS)*, pages 1–5, 2019. doi: 10.1109/WFCS.2019.8758024.

[55] Dharmini Shreenivas, Shahid Raza, and Thiemo Voigt. Intrusion detection in the rpl-connected 6lowpan networks. In *Proceedings of the 3rd ACM International Workshop on IoT Privacy, Trust, and Security*, page 31–38. Association for Computing Machinery, 2017. ISBN 9781450349697. doi: 10.1145/3055245.3055252.

[56] R. Doshi, N. Apthorpe, and N. Feamster. Machine learning ddos detection for consumer internet of things devices. In *2018 IEEE Security and Privacy Workshops (SPW)*, pages 29–35, 2018. doi: 10.1109/SPW.2018.00013.

[57] Liqun Liu, Bing Xu, Xiaoping Zhang, and Xianjun Wu. An intrusion detection method for internet of things based on suppressed fuzzy clustering. *EURASIP Journal on Wireless Communications and Networking*, 2018(1):113, 2018. ISSN 1687-1499. doi: 10.1186/s13638-018-1128-z.

[58] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani. A detailed analysis of the kdd cup 99 data set. In *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, pages 1–6, 2009.

[59] C. Cervantes, D. Poplade, M. Nogueira, and A. Santos. Detection of sinkhole attacks for supporting secure routing on 6lowpan for internet of things. In *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, pages 606–611, 2015. doi: 10.1109/INM.2015.7140344.

[60] N. Moustafa and J. Slay. Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set). In *2015 Military Communications and Information Systems Conference (MilCIS)*, pages 1–6, 2015.

[61] Canadian Institute of Cybersecurity. Nsl-kdd dataset, 2009. URL https://www.unb.ca/cic/datasets/nsl.html.

[62] Shahid Raza, Linus Wallgren, and Thiemo Voigt. Svelte: Real-time intrusion detection in the internet of things. *Ad Hoc Networks*, 11(8):2661 – 2674, 2013. ISSN 1570-8705. doi: https://doi.org/10.1016/j.adhoc.2013.04.014.

[63] A. Meddeb. Internet of things standards: who stands out from the crowd? *IEEE Communications Magazine*, 54(7):40–47, 2016. ISSN 1558-1896. doi: 10.1109/MCOM.2016.7514162.

[64] Huiping Guo, Yingjiu Li, and Sushil Jajodia. Chaining watermarks for detecting malicious modifications to streaming data. *Information Sciences*, 177(1):281 – 298, 2007. ISSN 0020-0255. doi: https://doi.org/10.1016/j.ins.2006.03.014. URL http://www.sciencedirect.com/science/article/pii/S0020025506000855.

[65] Wei Zhang, Yonghe Liu, Sajal K. Das, and Pradip De. Secure data aggregation in wireless sensor networks: A watermark based authentication supportive approach. *Pervasive and Mobile Computing*, 4(5):658 – 680, 2008. ISSN 1574-1192. doi: https://doi.org/10.1016/j.pmcj.2008.05.005. URL http://www.sciencedirect.com/science/article/pii/S157411920800059X.

[66] Sean Weerakkody, Omur Ozel, Yilin Mo, and Bruno Sinopoli. Resilient control in cyber-physical systems: Countering uncertainty, constraints, and adversarial behavior. *Foundations and Trends® in Systems and Control*, 7(1-2):1–252, 2019. ISSN 2325-6818. doi: 10.1561/2600000018. URL http://dx.doi.org/10.1561/2600000018.

[67] Khizar Hameed, Abid Khan, Mansoor Ahmed, Alavalapati Goutham Reddy, and M. Mazhar Rathore. Towards a formally verified zero watermarking scheme for data integrity in the internet of things based-wireless sensor networks. *Future Generation Computer Systems*, 82:274 – 289, 2018. ISSN 0167-739X. doi: https://doi.org/10.1016/j.future.2017.12.009.

[68] Zhang Guoyin, Kou Liang, Zhang Liguo, Liu Chao, Da Qingan, and Sun Jianguo. A new digital watermarking method for data integrity protection in the perception layer of iot. *Security and Communication Networks*, 2017, 2017. ISSN 1939-0114. doi: https://doi.org/10.1155/2017/3126010.

[69] Arwa Alromih, Mznah Al-Rodhaan, and Yuan Tian. A randomized watermarking technique for detecting malicious data injection attacks in heterogeneous wireless sensor networks for internet of things applications. *Sensors (Basel, Switzerland)*, 18(12):4346, Dec 2018. ISSN 1424-8220. doi: 10.3390/s18124346. URL https://pubmed.ncbi.nlm.nih.gov/30544877.

[70] Wei Li Neal N. Xiong Baowei Wang, Weiwen Kong. A dual-chaining watermark scheme for data integrity protection in internet of things. *Computers, Materials & Continua*, 58(3):679–695, 2019. ISSN 1546-2226. doi: 10.32604/cmc.2019.06106. URL http://www.techscience.com/cmc/v58n3/23040.

[71] Xi Shi and Di Xiao. A reversible watermarking authentication scheme for wireless sensor networks. *Inf. Sci.*, 240:173–183, August 2013. ISSN 0020-0255. doi: 10.1016/j.ins.2013.03.031. URL https://doi.org/10.1016/j.ins.2013.03.031.

[72] Qun Ding, Baowei Wang, Xingming Sun, Jinwei Wang, and Jian Shen. A reversible watermarking scheme based on difference expansion for wireless sensor networks. *International Journal of Grid Distribution Computing*, 8(2):143–154, 2015. doi: 10.14257/ijgdc.2015.8.2.14. URL http://dx.doi.org/10.14257/ijgdc.2015.8.2.14.

[73] Xingming Sun, Jianwei Su, Baowei Wang, and Qi Liu. Digital watermarking method for data integrity protection in wireless sensor networks. *International Journal of Security and Its Applications*, 7(4):407–416, 2013.